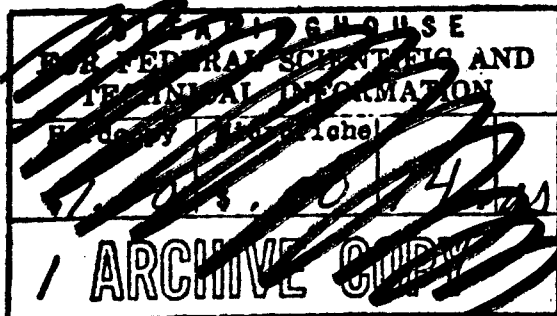AFOSR 66-1566

Final Report

to

Directorate of Information Sciences

Office of Aerospace Research

Air Force Office of Scientific Research

Rosslyn, Virginia 22209

Research

on

Grammatical Analysis

of

Spoken Language

Grant AF-AFOSR-22-64

February 2, 1964 – February 1, 1966

The University of Michigan

Ann Arbor

July 1966

June E. Shoup
Communication Sciences Laborator

1. Introduction

This is a report of research which has been conducted with the Directorate of Information Sciences of the Air Force Office of Scientific Research on the Grammatical Analysis of Spoken Language. This research project began in October, 1962, under Grant AF-AFOSR-22-63. Beginning in February, 1964, the research was continued under Grant AF-AFOSR-62-64. The research under this latter grant was carried on through January, 1965, and was subsequently extended through January, 1966.

The basic objective of the project has been to contribute to an increased understanding of the grammar of spoken language. The machine translation of languages is the chief applied motivation for much of the formal work that has been done on the grammar of natural language during recent years. Machine translation has thus far been primarily concerned with printed language. Automatic speech recognition, however, requires a syntax for spoken language, which may differ in several important aspects from the syntax of printed language.

Certain of the distinctions between spoken and written language are relatively obvious. For example, printed language usually consists of linear strings of discrete symbols, whereas spoken language consists of continuous acoustic waves. Since none of the major aspects of printed

or of spoken language are fully understood, however, it is
not surprising that it is exceedingly difficult to formali
the distinctions between these two forms of natural langua

2. Research Program

In spoken language several different types of infor
mation are transmitted concurrently. Information is trans
mitted by the phonemic sequences and simultaneously by the
associated prosodies. In spoken language there is a great
variation in syntax, from single lexical items to long and
involved expressions in which the relationships among the
lexical items may be obscure. In spoken language the mood
of the speaker, e.g. his attitude toward what he is saying
may vary from one utterance to another, and these differen
are often expressed in very subtle ways in the prosodies
of the dialect. It was not possible to investigate such
distinctions within the scope of the present project, but
has been possible to identify a number of the basic proble
which must be resolved before a convincing grammar of spok
language can be constructed. The research which has been
conducted is divided into four major parts, as discussed
below.

A. Prosodic Analysis. Techniques for research on
spoken language are somewhat restricted, and each has
major limitations. There are two primary techniques which

have been employed in the past. The first technique is that
of speech analysis and uses utterances from natural speech
as test signals. We are at present very far from being
able to make adequate instrumental interpretations of
utterances as they occur in natural conversational speech.
As a result, it is customary to conduct tests with utterances
which differ in some minimal linguistic way, and then to
study the physiological or the acoustical bases for these
differences. While much important information has been
obtained by this technique, it is exceedingly difficult to
obtain an integrated understanding of spoken language from
the accumulation of such information.

The second major technique is that of speech
synthesis. In speech synthesis it is possible to control
the analog of physiological variables or to control the
acoustical variables in a systematic way. The resulting
utterances may then be subjected to listening tests for
interpretation. Speech synthesis has the very great
advantage that the variables can be controlled systematically
and according to specification. It has the disadvantage
that there is no assurance that the synthesizer generates
the most relevant physiological or acoustical parameters.

A third technique was developed under the present
project for research on spoken language, particularly the
prosodies. This technique employs selective distortion of

the speech signal so that certain types of information are obliterated. With this technique, the phonemic information may be reduced effectively to zero, and only prosodic information retained. In addition, fundamental voice frequency may optionally be included or excluded. Essentially, the experimental system flattens the power spectrum without changing the instantaneous power, and then optionally reintroduces harmonics of the fundamental frequency. Thus the acoustic prosodic parameters of average fundamental voice frequency, average speech power, and acoustic phonetic duration may be preserved. The phonetic quality of vowels and consonants, including information about the secondary phonetic parameters is destroyed. A simple extension of the technique reported would also make it possible optional to include or exclude variations in average speech power.

The procedure provides a means of investigating the information contributed by the prosodies, either singly or jointly. A system to perform the above indicated distortions was constructed, and an experiment with the system was conducted. It was found that the system did, indeed, obliterate the phonemic information, while the prosodic information was retained. When the channel for average fundamental voice frequency was eliminated, correct listener responses to stress on English words decreased only slightly but the listener responses to intonation approached the chance level.

It is well known that the contributions to speech intelligibility of various frequency bands throughout the frequency domain do not summate linearly. It seems reasonable to assume that the contributions to utterance intelligibility of the phoneme sequences and of the prosodies singly and in combination also will not summate linearly. While it was not possible to investigate this subject during the course of the research, the technique described should provide the basis for investigating this aspect of speech intelligibility.

B. <u>Lexical Units</u>. In natural language the concept of lexical unit must be interpreted in a very broad sense. In the case of spoken language, the elemental meaningful units are denoted both by phoneme sequences and by prosodemes. The morpheme has been considered the basic meaningful unit of grammar in this study. An attempt was first made to specify an orthographic morpheme for a graphemic system of writing. A language may be written with graphemes, with a syllabary, or with ideographs, but the specification of the orthographic morpheme which has been developed is restricted to sequences of graphemes. An initial formulation of the morpheme in spoken language has also been constructed. The work which has been done is only preliminary and is not at present ready for publication. The principal investigator expects to continue work in this

area, however, particularly on the concept of the morpheme in spoken language.

C. Format for Syntactic Description. Because of the tremendous complexity of natural language, the development of a suitable format for expressing the structure of the syntax of spoken language presents a major problem. We have found previously that expressing the phonology of a dialect entirely in the form of rules results in an intricate and relatively obscure description. The format can be made more simple, more convenient, and easier to interpret by the use of a system of reference tables to which more general rules refer.

It seemed reasonable that this method might also be of value at the syntactic level. Reference tables make it possible to specify allowable and excluded sequences in a relatively convenient and direct manner. Since French was the native language of the investigator for this part of the study, the generation of verbal forms in French was selected as the topic. A generative format was employed and it was found that a relatively simple set of rules could be employed to refer to a set of tables for generating the various French verbs. There are, of course, many irregular forms and many forms which do not occur. The tables make it possible to express all of these various conditions in a relatively complete and compact manner.

While the above indicated investigation was concerned primarily with a very small part of the total syntax of the language, it offers an approach which seems worthy of much further consideration. Ultimately it should be possible to formulate such grammatical descriptions in either an analytic or a generative form, that is for either the analysis of utterances in a dialect or for generating utterances in a dialect.

D. **English Grammar.** Most of the research which has been done on automatic speech recognition has been concerned with English. As a result, English grammar is of particular interest to those working in the field of speech automation. During the course of the present project consideration has been given to the most appropriate form for a grammar for automatic speech recognition. The construction of a total grammar is a problem which obviously extends far beyond the scope of the present study. As a result a relatively small problem in English grammar was selected for investigation. The problem was further simplified by disregarding the phonological, particularly the prosodic, aspect of the grammar. Thus it was formulated primarily in terms of orthography.

Determiner phrases in English were chosen as the subject of investigation. The study began with an attempt to write a complete generative statement for a selected

set of determiner phrases. Many difficulties arose in the formulation. It proved to be extremely difficult to generate all reasonable forms and to exclude all unreasonable forms. This circumstance lead to a detailed consideration of the place of semantics in syntactic descriptions and of the properties required of a syntactic description. The study has suggested that the syntactic part of a grammatical description may be less important than the semantic part. The study has emphasized the importance of finding a practical and effective way to manage semantic data in grammatical descriptions.

3. Personnel

The following students have been employed on a part-time basis on the project with the Directorate of Information Sciences during the past two year period of the grant.

> Andre-Pierre Benguerel, Communication Sciences
> Ralph H. Fertig, Mathematics
> John R. Hanne, Communication Sciences
> George L. Huttar, Linguistics
> James A. Mason, Communication Sciences

4. Publications

Most of the areas discussed above involve basic problems which require continued investigation. Within the course of the project, however, certain particular studies were completed and were prepared for publication.

Manuscripts describing work carried out on the project

with the Information Sciences Directorate are as follows:

> Michael H. O'Malley and Gordon E. Peterson,
> An Experimental Method for Prosodic Analysis,
> Phonetica, (accepted for publication).

Some of the more difficult questions in the study of language involve the nature and function of the prosodies. While the prosodies have been investigated by observing their acoustic correlates and by varying the relevant acoustical parameters in synthetic speech, the use of distorted natural speech also provides an effective procedure for perceptual studies. In this study a technique for reducing a speech wave to the acoustic-prosodic parameters of speech power, phonetic duration, and fundamental voice frequency was developed. With the system described, all suprasegmental information based on the prosodic parameters is transmitted while all segmental information is destroyed. The technique consists of multiplying the input by a random telegraph wave, thus flattening the power spectrum. Harmonics of the fundamental frequency are then optionally reintroduced to provide fundamental frequency information. Listening tests showed that phoneme intelligibility was almost eliminated while intonation and stress were only slightly affected. Furthermore, eliminating the fundamental voice frequency caused the perception of intonation to approach the chance level while the perception of stress was only slightly affected. The technique should be useful for investigating the role of the prosodies in grammatical structures.

> Andre-Pierre Benguerel, Generation of Verbal
> Forms in French, The International Journal of
> American Linguistics (submitted for publication).

This paper presents a generative grammar of French verbal forms. It consists of an ordered set of rewrite rules and of a set of tables. It generates all existing verbal forms without generating any nonexisting ones. To shorten the part of the grammar employing rewrite rules, symbols with indices have been used. The tables present stem and ending distributions in matrix form and the indices of the complex symbols correspond to the different row and column headings of these matrices. If the values of the indices are chosen in such a way as to correspond to a nonexisting form, i.e. to an empty entry in a matrix, the whole string is deleted and no incorrect form can be generated.

The verbs are distributed into classes, according to their stem distribution, stem formation, and endings. A good compromise is reached between a large number of classes with few stems and a small number of classes with many stems, a large number of which would often be identical. The number of ending paradigms has also been kept at a minimum. Nevertheless the use of more than one ending paradigm has proven worthwhile in decreasing the number of verb classes. The presentation of the material in tabular form may appear to be lengthy, but actually it makes possible the presentation of a large amount of grammatical detail in a compact form. Although the total number of rules (or of choices to be made) may be larger than in a morphophonemic description, the description is more exact and the average number of choices per production is smaller. In other words, once we have selected a verb class, the other 62 classes are excluded and the number of choices that remain to be made is certainly smaller than the number of exceptions that would have to be looked up in a complete morphophonemic description.

The departure from an ordinary generative grammar lies in the use of a tabular form for presenting the lexical material. This return to a presentation often found in traditional grammars has several advantages: 1) It makes it more readable to anyone who wants to follow through the generation of a verb form. 2) It is readily usable in a computer program. 3) It naturally complements complex symbols, since row and column headings of the matrices are actual realizations of the indices. 4) It can be used as a teaching tool with little modification. 5) It is probably closer to the intuition of the literate native speaker than a system consisting purely of rewrite rules, without any device such as complex symbols or tables.

James A. Mason and Gordon E. Peterson, On the Problem of Describing the Grammar of Natural Languages, Language and Speech, (accepted for publication).

Difficulties encountered in an attempt to describe the syntax of English determiner phrases resulted in a reconsideration of the purpose and organizational principles of grammatical descriptions. Some illumination of the nature of grammatical descriptions is obtained by a consideration of systems of chess notations. Problems of grammatical description discussed with reference to two specific examples of chess notations include:

(1) The problem of describing the "basic regularities" which determine how the sentences of a language are understood;

(2) The problem of translating between two languages with the same "universe of discourse" but with different ways of referring to it;

(3) The problem of explaining the intuitive notion of "grammaticality" which native users of a language possess.

The importance, for a useful language description, of describing the "semantic interpretation" process is illustrated and emphasized, and the value of language descriptions in the form of generative grammars is questioned.

DOCUMENT CONTROL DATA — R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY *(Corporate author)*	2A. REPORT SECURITY CLASSIFICATION
Communications Sciences Lab. University of Michigan Ann Arbor, Mich. 48105	☒ Unclassified Other — Specify 2B. GROUP

3 REPORT TITLE

RESEARCH ON GRAMMATICAL ANALYSIS OF SPOKEN LANGUAGE

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

☐ Scientific Report ☒ Final Report ☐ Journal Article ☐ Proceedings ☐ Book

5. AUTHOR(S) *(Last name, first name, initial)*

SHOUP, JUNE E.

6. REPORT DATE AS PRINTED	7A. TOTAL NO. OF PAGES	7B. NO. OF REFS
July 66	11	3

8A CONTRACT OR GRANT NO. AF-AFOSR-22-64	9A. ORIGINATOR'S REPORT NUMBER(S) *(if given)*
B. PROJECT NO. 9769	
C. 61445014	9B. OTHER REPORT NO.(S) *(Any other numbers that may be assigned this report)* AFOSR 66-1566
D.	AD

10. AVAILABILITY/LIMITATION NOTICES	
Distribution of this document is unlimited.	☒ Available from DDC ☒ Available from CFSTI ☐ Available from Source ☐ Available Commercially

11. SUPPLEMENTARY NOTES *(Citation)*	12. SPONSORING MILITARY ACTIVITY
	AF Office of Scientific Research (SRI) Office of Aerospace Research Washington, D. C. 20333

13 ABSTRACT

This report summarizes research directed toward an increased understanding of the grammar of spoken language. A technique for prosodic analysis was developed that flattens the power spectrum without changing the instantaneous power and optionally reintroduces harmonics of the fundamental frequency. When the channel for average fundamental voice frequency was eliminated, correct listener responses to stress on English words decreased only slightly, but listener responses to intonation approached the chance level. A generative format was developed for expressing the structure of the syntax of a spoken language (French). An ordered set of rewrite rules used with a set of tables generated all existing verb forms and no nonexisting ones. An attempt to describe the syntax of English determiner phrases led to a reconsideration of the place of semantics in syntactic descriptions and the properties required of a syntactic description. Problems of grammatic description were examined with reference to two examples of chess notations. Abstracts of three papers written during the period are included.

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Language analysis						
Speech						
Syntactic analysis						
Semantics						
Grammar						
Prosodic analysis						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

_____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

_____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

_____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.